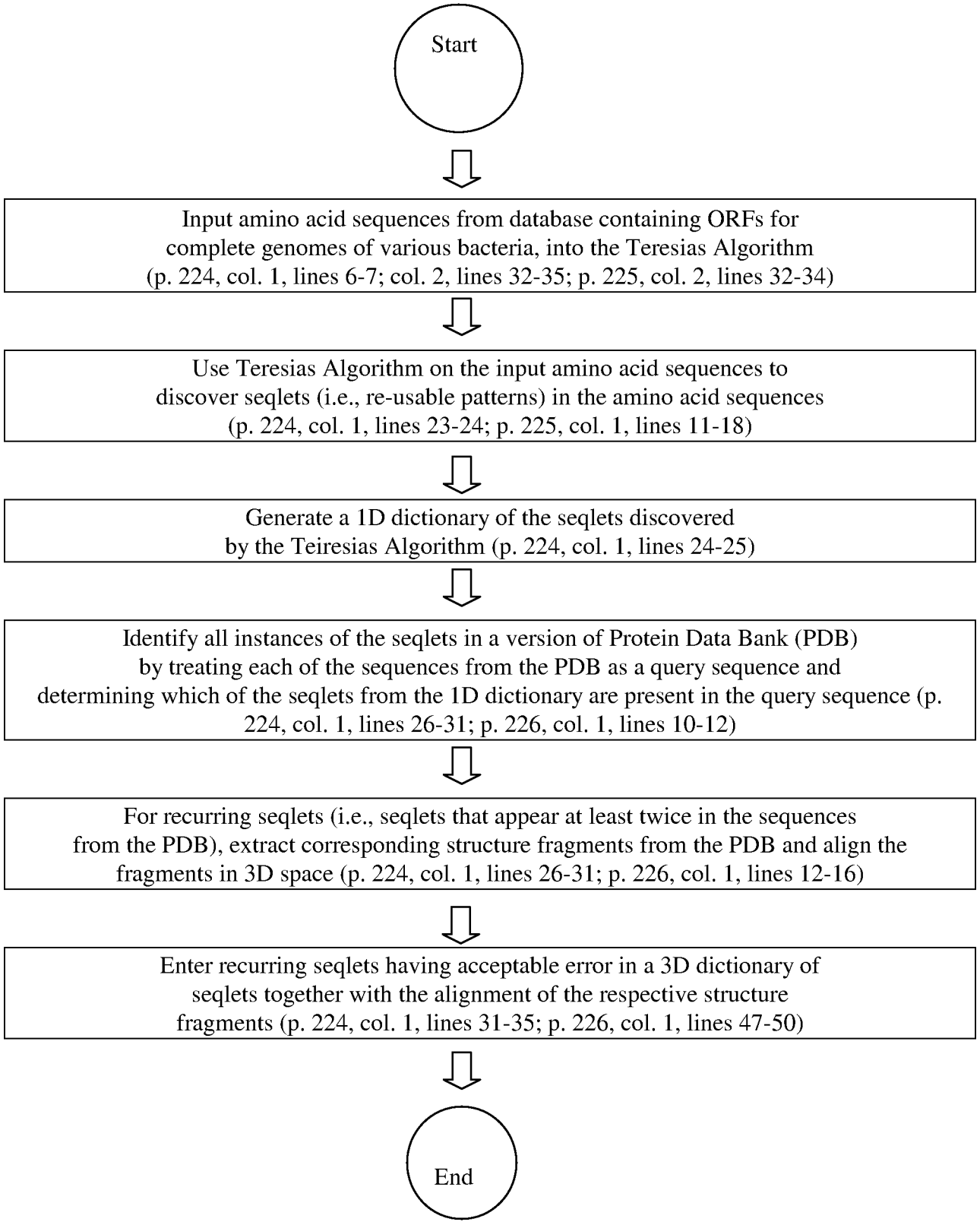
Start
Input amino acid sequences from database containing ORFs for complete genomes of various bacteria, into the Teresias Algorithm (p. 224, col. 1, lines 6-7; col. 2, lines 32-35; p. 225, col. 2, lines 32-34)
Use Teresias Algorithm on the input amino acid sequences to discover seqlets (i.e., re-usable patterns) in the amino acid sequences (p. 224, col. 1, lines 23-24; p. 225, col. 1, lines 11-18)
Generate a 1D dictionary of the seqlets discovered by the Teiresias Algorithm (p. 224, col. 1, lines 24-25)
Identify all instances of the seqlets in a version of Protein Data Bank (PDB) by treating each of the sequences from the PDB as a query sequence and determining which of the seqlets from the 1D dictionary are present in the query sequence (p. 224, col. 1, lines 26-31; p. 226, col. 1, lines 10-12)
For recurring seqlets (i.e., seqlets that appear at least twice in the sequences from the PDB), extract corresponding structure fragments from the PDB and align the fragments in 3D space (p. 224, col. 1, lines 26-31; p. 226, col. 1, lines 12-16)
Enter recurring seqlets having acceptable error in a 3D dictionary of seqlets together with the alignment of the respective structure fragments (p. 224, col. 1, lines 31-35; p. 226, col. 1, lines 47-50)
End